

Query Dependent Prompt System

Prompt Content

- Prompt 18:

First, assess if the news topic could interest a wide Persian-speaking audience. Then, determine if it aligns with significant themes such as major economic, political, or social events. If the news is relevant to many people, it is 'important' (1); if it targets a specific or small group, it is 'not important' (0).

Conclude your analysis of these scenarios by presenting the final classification in the following format:

Final Classification: [1 or 0]

Prompt Content

- Prompt 18:

First, assess if the news topic could interest a wide Persian-speaking audience. Then, determine if it aligns with significant themes such as major economic, political, or social events. If the news is relevant to many people, it is 'important' (1); if it targets a specific or small group, it is 'not important' (0).

Conclude your analysis of these scenarios by presenting the final classification in the following format:

Final Classification: [1 or 0]

Prompt Content

- Prompt 19: (Chain-of-thought)

The goal is to classify news items into 'important' (1) or 'not important' (0). To classify accurately, follow these steps:

1. Identify if the news topic could be relevant to a large Persian-speaking audience.

2. Assess if it pertains to significant economic events (currency or inflation changes, housing updates, etc.), critical political events (government actions, international relations), or socially impactful themes that could affect many people.

3. Finally, determine if the story has widespread appeal or is only relevant to a niche audience.

If the news is of broad interest and covers the themes above, label it as '1'. Otherwise, label it as '0'. Upon completing the scenario analysis, output the final classification using the format below:

Final Classification: [1 or 0]

Prompt Content

- Prompt 21: (Tree-of-thought)

For this classification task, follow these branches of thought to decide if the news item is 'important' (1) or 'not important' (0):

1. First, examine if the topic could affect a large portion of Persian-speaking users, considering its potential reach.

2. Next, break down the topic's content into economic, political, and social relevance.

- For economic news: Consider factors like inflation, housing, and stock trends relevant to everyday users.

- For political news: Evaluate if the content relates to Iran's major policies, high-profile government changes, or global interactions.

- For social relevance: Check if it covers popular sports or events with a broad appeal.

3. Assess if the story's appeal is universal or niche.

Assign '1' if the story is broadly significant or '0' if it appeals mainly to a niche audience. Once you've reviewed these scenarios, provide the final classification formatted as follows:

Final Classification: [1 or 0]

Prompt Content

- Prompt 24: (Chain-of-thought)

هدف این است که اخبار را به دو دسته «مهم» (۱) و «غیر مهم» (۰) طبقه‌بندی کنیم. برای طبقه‌بندی دقیق، مراحل زیر را دنبال کنید:

۱. بررسی کنید که آیا موضوع خبر می‌تواند برای جمعیت زیادی از فارسی‌زبانان مرتبط باشد یا خیر.

۲. ارزیابی کنید که آیا خبر به رویدادهای مهم اقتصادی (تغییرات ارز یا تورم، به‌روزرسانی‌های مسکن و غیره)، رویدادهای سیاسی حیاتی (اقدامات دولت، روابط بین‌الملل) یا موضوعات اجتماعی تاثیرگذار که ممکن است افراد زیادی را تحت تاثیر قرار دهد، مربوط می‌شود.

۳. در نهایت، مشخص کنید که آیا خبر جذابیت عمومی دارد یا تنها برای مخاطبان خاصی قابل توجه است.

اگر خبر به موضوعات ذکر شده مربوط باشد و علاقه‌مندی گسترده‌ای را به خود جلب کند، آن را با «۱» برچسب بزنید. در غیر این صورت، «۰» را انتخاب کنید. پس از بررسی این سناریوها، طبقه‌بندی نهایی را در قالب زیر ارائه کنید:

طبقه‌بندی نهایی: «۰ یا ۱»

Prompt Content

- Prompt 25: (Tree-of-thought)

برای این وظیفه‌ی طبقه‌بندی، از شاخه‌های فکری زیر استفاده کنید تا تصمیم بگیرید که آیا خبر «مهم» (۱) است یا «غیر مهم» (۰):

۱. ابتدا بررسی کنید که آیا موضوع خبر می‌تواند بخش بزرگی از کاربران فارسی‌زبان را تحت تأثیر قرار دهد و به گستردگی احتمالی آن توجه کنید.
۲. سپس محتوای موضوع را از نظر اهمیت اقتصادی، سیاسی و اجتماعی تحلیل کنید.

برای اخبار اقتصادی: عواملی مانند تورم، وضعیت مسکن و روندهای بورس که برای کاربران عادی اهمیت دارند را مدنظر قرار دهید.

برای اخبار سیاسی: ارزیابی کنید که آیا محتوا به سیاست‌های کلان ایران، تغییرات مهم در دولت، یا تعاملات جهانی مربوط می‌شود.

برای اهمیت اجتماعی: بررسی کنید که آیا خبر شامل رویدادهای ورزشی محبوب یا موضوعاتی با جذابیت گسترده است.

۳. بررسی کنید که آیا جذابیت خبر عمومی است یا فقط برای مخاطبان خاصی جذابیت دارد.

اگر خبر از نظر گسترده‌ای مهم است، آن را با «۱» برچسب بزنید. در غیر این صورت، اگر بیشتر برای مخاطبان خاص جذاب است، «۰» را انتخاب کنید. تحلیل این سناریوها را به پایان رسانده و طبقه‌بندی نهایی را به صورت زیر ارائه دهید:

طبقه بندی نهایی: «۰ یا ۱»

Zero-shot Train Dataset Results

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Prompt 18	54%	57%	60%	51%	40%	63%
Prompt 19	68%	59%	62%	59%	40%	78%
Prompt 21	48%	57%	59%	47%	40%	55%
Prompt 24	74%	58%	56%	57%	29%	84%
Prompt 25	71%	56%	56%	56%	31%	82%

Zero-shot Train Dataset Results

Title Only	Sports	Political	Economic	Global	Social	Cultural	Scientific	Health
Prompt 18	56%	43%	59%	43%	49%	41%	49%	52%
Prompt 19	59%	54%	62%	48%	57%	49%	54%	58%
Prompt 21	52%	41%	51%	40%	44%	37%	47%	49%
Prompt 24	55%	56%	54%	56%	55%	48%	52%	49%
Prompt 25	55%	54%	54%	56%	58%	49%	58%	49%

Zero-shot Eval Dataset Results

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Prompt 18	51%	56%	59%	48%	36%	61%
Prompt 19	66%	57%	60%	56%	36%	77%
Prompt 21	45%	56%	58%	44%	36%	52%
Prompt 24	76%	58%	57%	57%	30%	85%
Prompt 25	72%	57%	57%	57%	31%	83%

Zero-shot Eval Dataset Results

Title Only	Sports	Social	Political	Health	Cultural	Scientific	Economic	Global
Prompt 18	47%	47%	42%	20%	46%	49%	54%	42%
Prompt 19	52%	59%	52%	23%	61%	47%	57%	46%
Prompt 21	42%	41%	39%	29%	44%	42%	47%	38%
Prompt 24	57%	62%	56%	37%	47%	57%	54%	49%
Prompt 25	58%	57%	57%	28%	52%	58%	55%	47%

Zero-shot Test Dataset Results

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Prompt 18	53%	57%	63%	49%	35%	63%
Prompt 19	72%	62%	68%	63%	44%	81%
Prompt 21	46%	57%	61%	44%	34%	54%
Prompt 24	77%	55%	54%	54%	22%	86%
Prompt 25	74%	54%	54%	54%	24%	84%

Zero-shot Test Dataset Results

Title Only	Social	Sports	Political	Cultural	Global	Economic	Scientific	Health
Prompt 18	45%	55%	40%	41%	35%	59%	58%	63%
Prompt 19	56%	61%	49%	49%	41%	62%	58%	80%
Prompt 21	45%	49%	35%	35%	34%	48%	49%	63%
Prompt 24	48%	50%	58%	47%	46%	51%	56%	69%
Prompt 25	54%	51%	54%	60%	43%	52%	46%	69%

Zero-shot Query Dependent Results

- Query Dependent Prompt System Result:

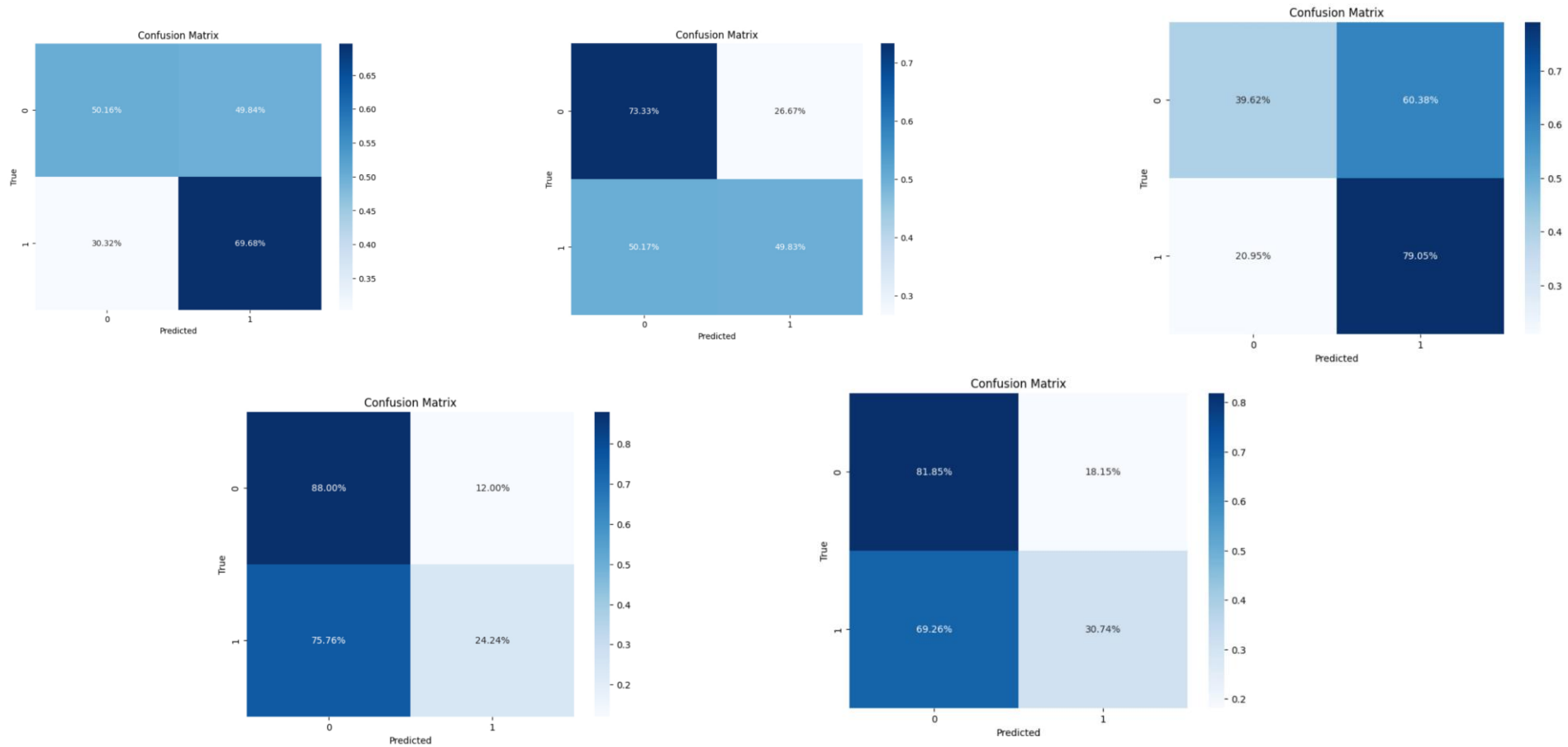
Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Q-Dependent	86%	75%	72%	73%	55%	92%

- Compared to best zero-shot prompt results:

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Prompt 19	72%	62%	68%	63%	44%	81%

Zero-shot Query Dependent Results

- This is because of different behaviors of prompts in terms of confusion matrix:



Few-shot Test Dataset Results

- All of these results conducted in k=20 setting.

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Prompt 18	66%	62%	71%	60%	43%	76%
Prompt 19	76%	64%	72%	65%	47%	83%
Prompt 21	70%	63%	72%	62%	45%	79%
Prompt 24	82%	68%	70%	69%	49%	89%
Prompt 25	82%	68%	69%	69%	48%	89%

Few-shot Test Dataset Results

- All of these results conducted in k=20 setting.

Title Only	Social	Sports	Political	Cultural	Global	Economic	Scientific	Health
Prompt 18	66%	57%	48%	61%	48%	64%	58%	71%
Prompt 19	65%	62%	56%	76%	55%	69%	62%	89%
Prompt 21	63%	56%	53%	67%	52%	67%	60%	80%
Prompt 24	66%	68%	66%	68%	68%	70%	57%	42%
Prompt 25	69%	65%	66%	66%	63%	72%	50%	68%

Few-shot Test Dataset Results

- Query Dependent Prompt System Result on zero-shot setting:

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Q-Dependent	86%	75%	72%	73%	55%	92%

- Best few-shot prompt results:

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Prompt 24	82%	68%	70%	69%	49%	89%

- Best zero-shot prompt results:

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Prompt 19	72%	62%	68%	63%	44%	81%

Few-shot Train Dataset Results

- All of these results conducted in k=20 setting.

Title Only	Accuracy	Precision	Recall	F1-Score	F1-Score('1')	F1-Score('0')
Prompt 18	99%	97%	99%	98%	97%	99%
Prompt 19	99%	98%	99%	99%	98%	99%
Prompt 21						
Prompt 24						
Prompt 25						

- The problem here is that the target title is include in examples provided in few-shot prompt!